Point-to-Ellipse and Point-to-Ellipsoid Distance Equation Analysis☆

Alexei Yu. Uteshev*, Marina V. Goncharova

St. Petersburg State University, Universitetskij pr. 35, Petrodvoretc, St. Petersburg 198504, Russia

Abstract

For the problem of distance evaluation from a point X_0 to an ellipse in $\mathbb{R}^2$ and to an ellipsoid in $\mathbb{R}^3$ given by algebraic equation $G(X) = 0$, we investigate the properties of the *distance equation*, i.e. an algebraic equation whose zeros coincide with the critical values of the squared distance function. We detail the structure of this equation and an algorithm for finding the point in the quadric nearest to X_0. We also find analytical formulas for distance approximations using the expansion of the zero of distance equation into power series $\sum_{j=1}^{\infty} \ell_j G^j(X_0)$. Exact values for the approximation error bounds are obtained via construction of an analogue of the distance equations for the curve-to-curve distance problem.

Keywords: Distance to ellipse, distance to ellipsoid, distance approximation, projection of a point, Sampson's distance

1. Introduction

The problem of finding the value of the distance from a given point X_0 to a given quadric in $\mathbb{R}^n$ is of great importance for several branches of mathematics, statistical data analysis, astronomy, particle physics and image processing.

☆This document is a collaborative effort.

*Corresponding author

Email addresses: alexeiuteshev@gmail.com (Alexei Yu. Uteshev), marina.yashina@gmail.com (Marina V. Goncharova)

Preprint submitted to J. Computational and Applied Mathematics *July 3, 2017*

The current approaches for solving this problem can be categorised into either numerical or analytical methods. The first one consists in construction of an appropriate iterative procedure for solving the nonlinear constrained optimization problem: the aim is to generate a sequence of points in $\mathbb{R}^n$ converging to the point in the quadric nearest to X_0 (projection of X_0 onto the quadric) [1, 2, 3]. This might be effective for problems when the treated quadric is assumed to be *fixed*, i.e. when all of its coefficients are *specialized*. However, in several industrial applications, the parameters of the quadric may *variate*, such as, for instance, when it simulates an object moving in $\mathbb{R}^2$ or in $\mathbb{R}^3$ [4]. To solve the problem in such a statement, an analytical expression is needed as a function of parameters either for the distance or for its suitable approximation.

The necessity for the analytical (symbolical) representation also stems from the problem of the approximation of scattered data known as the *ellipse* (*ellipsoid*) *fitting problem* [5, 6, 7, 8]. The latter consists in finding the coefficients of the second order algebraic equation

$$G(X) := X^T \mathbf{A} X + 2\, B^T X + c = 0 \tag{1.1}$$

providing the ellipse (or ellipsoid) closest to the given set of test (measured) data points $\{X_j\}_{j=1}^N \subset \mathbb{R}^n$. The first obstacle in solving this problem consists in the evaluation of the closeness of a given point X_0 to the quadric (1.1) with *undetermined coefficients*. Since the explicit formula for the distance function is unavailable, this closeness is usually evaluated by simpler computed substitutes, like, for instance, the *algebraic distance* $|G(X_0)|$ (or $G^2(X_j)$) or *Sampson's distance*

$$d_S := \frac{|G(X_0)|}{2\sqrt{(\mathbf{A}X_0 + B)^T(\mathbf{A}X_0 + B)}} \, . \tag{1.2}$$

The comparison of effectiveness of this and some other substitution has been evaluated only empirically [9, 10, 11] since, for the reliable error analysis, the explicit expression for the distance function is needed.

The authors of the present paper have succeeded in finding the general expression for this *geometric*, i.e. Euclidean, *distance* from the point to the quadric (1.1) in $\mathbb{R}^n$ [12, 13]. The result has been achieved via the application of analytical (algebraic) methods of *elimination of variables* from the system of equations providing the coordinates of stationary points of the Lagrangian. The value of the squared distance is among the positive zeros of an appropriate univariate algebraic *distance equation* of the degree generically equal to $2n$ and with coefficients depending polynomially on those of the quadric (1.1) and the coordinates of X_0. In the framework of this approach, the coordinates of the point in the quadric nearest to X_0 can also be expressed in the form of an appropriate rational functions of a zero of the distance equation and parameters of the problem.

However, there exists an evident drawback in the obtained analytical expression for the distance function which blocks its immediate utilization for the applied problems mentioned above. It happens to be *implicit*.

In Conclusion section of the paper [13], one of the problems listed for further investigation is stated as follows:

"*To resolve the distance equation, i.e. to obtain approximate (but explicit) formulas for its zeros.*"

The aim of the present paper is to extract from the distance equation the explicit formulas for the distance function approximation and to estimate the tolerances for these approximations. Although some of the results considered below are valid for the general case of $\mathbb{R}^n$, we will be focused mainly onto the problems in $\mathbb{R}^2$ and $\mathbb{R}^3$.

The paper is organized as follows. Section 2 contains some preliminary results from the classical Elimination Theory including definition and some properties of the *discriminant* of the univariate polynomial and the *resultant* of a pair of such polynomials. The necessity in these notions is justified by the fact that the distance equation can be constructed via the appropriate discriminant computation. This is detailed in Section 3 where the complete symbolic expression for the distance equation for the point-to-ellipse problem is presented for the case of an ellipse given in canonical form. As for the point-to-ellipsoid distance equation, its representation is limited to just a few terms which will be of use in foregoing sections. We are also concerned here with the procedure of finding the coordinates of the nearest point in the considered quadrics.

In Section 4 we discuss some approaches for approximation of the zero of the distance equation. We first treat the one suggested in [14] consisting in reduction of this equation to its linearization in a vicinity of the zero value of the variable. We demonstrate that the validity of this approximation depends crucially on the assumption that the minimal positive zero of the distance equation equals the squared point-to-quadric distance. This assumption while being true generically, fails for some locations of the point X_0 in the major axis of the considered ellipse (or ellipsoid). We suggest an alternative approach for the approximation of the zero of distance equation.

We construct its expansion in the power series of the expression for $G(X_0)$ which can be treated as a small parameter in a vicinity of the considered quadric. The coefficients of this expansion

$$\ell_1 G(X_0) + \ell_2 G^2(X_0) + \ell_3 G^3(X_0) + \dots$$

are determined with the aid of expressions for the coefficients of distance equations obtained in Section 3. It turns out that the first approximation of the distance function obtained in this way coincides with Sampson's distance (1.2). We find the second approximation in the form

$$d_{(2)} = d_S \sqrt{1 + \frac{1}{2} \cdot \frac{(\mathbf{A}X_0 + B)^T \mathbf{A} (\mathbf{A}X_0 + B)}{[(\mathbf{A}X_0 + B)^T (\mathbf{A}X_0 + B)]^2} G(X_0)} \qquad (1.3)$$

and our next aim consists in comparison of the tolerances of all of the considered approximations.

This is tackled in Section 5. Given an expression $d_A(X_0)$ for the distance approximation from a point X_0 to the given ellipse or ellipsoid, we are interested in the geometry of the set of manifolds $d_A^2(X) = h$ in a vicinity of this quadric. For several approximation formulas in the planar case, the qualitative behavior of their *iso-value contours* have been compared in [9, 10, 11] . We intend to *verify the geometry with the aid of algebra*, i.e. we aim at evaluation of the exact bounds for displacement of the manifold $d_A(X) = h$ from the equidistant one to the approximated quadric. The underlying idea for the suggested approach is similar to that used in Section 3. For the points in the manifold $d_A^2 = h$ we treat the point-to-quadric distance function and look for its critical values. Reducing this problem of constrained optimization to that of solving of an appropriate algebraic equation system we then apply the resultant technique for computation of a univariate algebraic equation with the zero set coinciding with that of critical values of the distance function. We compute this equation for the point-to-ellipse distance approximations given by formulas (1.2) and (1.3). As for the Sampson's approximation in $\mathbb{R}^3$, we prove that the problem of finding the points of maximal possible error can be reduced to its counterpart in $\mathbb{R}^2$.

Some of the results of the paper have been presented at the Third International Conference on Applied and Computational Mathematics (Geneva, 2014). The result of Theorem 4.3 has been announced there as the one stemmed from the computer simulation. In the present paper we give a constructive proof for it. The results of Section 4 concerning the formal applicability of the suggested approximation formulas (for instance, nonnegativity of the radicand in (1.3)) have never been discussed; the results of Theorems 5.2, 5.3, 5.4 are also original.

Notation is kept to correlate with that from [12, 13]. We set $c = -1$ in (1.1), i.e. the ellipse or ellipsoid equation is considered in the form

$$G(X) := X^T \mathbf{A} X + 2\, B^T X - 1 = 0 \tag{1.4}$$

with X_0, B and X being the column vectors from $\mathbb{R}^n$. $\mathbf{A}$ is a symmetric sign-definite matrix of the order n, while $\mathbf{I}$ is the identity matrix of the same order. $\mathcal{D}_x$ denotes the discriminant of the polynomial (subscript denotes the variable w.r.t. which the polynomial is treated).

Accuracy. In the numerical examples we give the results of approximate computations rounded either to 10^{-5} or to 10^{-2}.

2. Algebraic Preliminaries

In the rest of the paper we utilize essentially some basic results from Elimination Theory (dealing with the algebraic algorithms of elimination the variables from a system of nonlinear algebraic equations), including the notions of the *resultant* and the **discriminant**. In the present section we will sketch only the results necessary for our particular purposes while for the systematic statement of the theory we refer to [15, 16].

The discriminant of the polynomial

$$F(\mu) = A_0 \mu^n + A_1 \mu^{n-1} + \ldots + A_n,\ A_0 \neq 0\,, \tag{2.1}$$

with the zeros $\mu_1, \ldots, \mu_n$ is formally defined as

$$\mathcal{D}_\mu(F(\mu)) := A_0^{2n-2} \prod_{1 \le j < k \le n} (\mu_k - \mu_j)^2\ .$$

Theorem 2.1. *Discriminant can be expressed in the form of a homogeneous polynomial function from the coefficients* $\{A_j\}_{j=0}^n$*:*

$$\mathcal{D}_\mu(F(\mu)) \equiv \mathfrak{D}(A_0, A_1, A_2, \ldots, A_n)\ .$$

Its vanishement is a necessary and sufficient condition for the existence of a multiple zero for $F(\mu)$*. Provided that this zero* μ_* *is unique and possesses the multiplicity* 2*, the following ratio is valid*

$$1 : \mu_* : \mu_*^2 : \ldots = \frac{\partial \mathfrak{D}}{\partial A_n} : \frac{\partial \mathfrak{D}}{\partial A_{n-1}} : \frac{\partial \mathfrak{D}}{\partial A_{n-2}} : \ldots \tag{2.2}$$

From the last statement follows that a unique double zero of a polynomial can expressed as a rational function of its coefficients.

For instance, the discriminant for the cubic polynomial can be computed by the formula

$$\mathcal{D}_{\mu}(A_0\mu^3 + A_1\mu^2 + A_2\mu + A_3) \tag{2.3}$$
$$= A_1^2A_2^2 - 4\,A_1^3A_3 - 4\,A_0A_2^3 + 18\,A_0A_1A_2A_3 - 27\,A_0^2A_3^2\,.$$

If it vanishes then the polynomial possesses a multiple zero, which, provided that its multiplicity equals 2, can be evaluated via the formula (2.2)

$$\mu_* = \frac{-9\,A_0A_3A_1 - A_1^2A_2 + 6\,A_2^2A_0}{27\,A_0^2A_3 - 9\,A_0A_1A_2 + 2\,A_1^3};$$

or, alternatively, via

$$\mu_* = \frac{9\,A_0A_3 - A_1A_2}{2(A_1^2 - 3\,A_0A_2)} \tag{2.4}$$

with the aid of determinantal representation for the discriminant [13].

The discriminant for the quartic polynomial can be computed as

$$\mathcal{D}_{\mu}(A_0\mu^4 + A_1\mu^3 + A_2\mu^2 + A_3\mu + A_4) = 4\,\mathfrak{I}_2^3 - 27\,\mathfrak{I}_3^2 \tag{2.5}$$

where

$$\begin{array}{rcl} \mathfrak{I}_2 & := & 4\,A_0A_4 - A_1A_3 + \frac{1}{3}A_2^2, \\ \mathfrak{I}_3 & := & -A_0A_3^2 - A_1^2A_4 + \frac{8}{3}A_0A_2A_4 + \frac{1}{3}A_1A_2A_3 - \frac{2}{27}A_2^3\,. \end{array}$$

If there exists a unique double zero for the polynomial then it can be evaluated via the formula

$$\mu_* = \frac{2\,A_1\mathfrak{I}_2^2 + (3\,A_1A_2 - 18\,A_0A_3)\mathfrak{I}_3}{(24\,A_0A_2 - 9\,A_1^2)\mathfrak{I}_3 - 8\,A_0\mathfrak{I}_2^2}\,. \tag{2.6}$$

Corollary 2.1. *If* $A_0 \neq 0, A_n \neq 0$ *then*

$$\mathcal{D}_{\mu}(A_0\mu^n + A_1\mu^{n-1} + \ldots + A_n) = \mathcal{D}_{\mu}(A_0 + A_1\mu + \ldots + A_n\mu^n)\,.$$

The notion of discriminant is a particular case of that of the **resultant** of a pair of univariate polynomials. For the polynomial $F(\mu)$ given by (2.1) and for the polynomial $G(\mu) = B_0^m + B_1^{m-1} + \ldots + B_m, B_0 \neq 0$ with the zeros $\eta_1, \ldots, \eta_m$, their resultant is formally defined as

$$\mathcal{R}_\mu(F, G) := A_0^m B_0^n \prod_{j=1}^n \prod_{k=1}^m (\mu_j - \eta_k) .$$

The validity of the counterpart for Theorem 2.1 can also be established. We refer to [15, 16] for the details (including the practical methods of computation of the resultant) and for application of this object to the problem of elimination of variable from the system of two bivariate algebraic equations.

3. Distance Equation

In [13] the general result has been presented for finding the distance from a point to an ellipsoid in $\mathbb{R}^n$:

Theorem 3.1. *The square of the distance to the ellipsoid (1.4) from the point X_0 not lying in it ($G(X_0) \neq 0$) coincides with the minimal positive zero t_* of the* ***distance equation***

$$\mathcal{F}(t) := \mathcal{D}_\mu(\Phi(\mu, t)) = 0 \tag{3.1}$$

provided that this zero is not a multiple one. Here

$$\Phi(\mu, t) := \det \left(\left[\begin{array}{cc} \mathbf{A} & B \\ B^T & -1 \end{array} \right] + \mu \left[\begin{array}{cc} -\mathbf{I} & X_0 \\ X_0^T & t - X_0^T X_0 \end{array} \right] \right) . \tag{3.2}$$

The coordinates of the point in the ellipsoid nearest to X_0 are as follows:

$$X_* = (\mu_* \mathbf{I} - \mathbf{A})^{-1} (B + \mu_* X_0) . \tag{3.3}$$

Here μ_* stands for the multiple zero of the polynomial $\Phi(\mu, t_*)$.

Once the minimal positive zero t_* of (3.1) is evaluated, the multiple zero μ_* for the polynomial $\Phi(\mu, t_*)$ can be expressed, with the aid of Theorem 2.1, as a rational function of t_* (with coefficients polynomially dependent on the coefficients of (1.1) and on coordinates of X_0).

We now intend to detail the above result for the particular cases of $\mathbb{R}^2$ and $\mathbb{R}^3$. This will be done by virtue of the characteristic polynomial of the matrix $\mathbf{A}$

$$f(\mu) := \det(\mu \mathbf{I} - \mathbf{A}) = \mu^n + c_1 \mu^{n-1} + \ldots + c_n .$$

With the aid of the Schur formula [17] for the determinant of a block matrix, one can convert the representation (3.4) for $\Phi(\mu,t)$ into the form

$$\Phi(\mu,t) \tag{3.4}$$

$$\equiv (-1)^n\left[f(\mu)(\mu(t-X_0^TX_0)-1)+(B+\mu X_0)^T\mathbf{adj}(\mu\mathbf{I}-\mathbf{A})(B+\mu X_0)\right].$$

Here $\mathbf{adj}(\mu\mathbf{I}-\mathbf{A})$ stands for the adjoint matrix to the matrix $\mu\mathbf{I}-\mathbf{A}$:

$$(\mu\mathbf{I}-\mathbf{A})\cdot\mathbf{adj}(\mu\mathbf{I}-\mathbf{A})\equiv f(\mu)\mathbf{I}\,;$$

and it can be computed as

$$\mathbf{adj}(\mu\mathbf{I}-\mathbf{A})$$

$$=\mathbf{A}^{n-1}+(\mu+c_1)\mathbf{A}^{n-2}+(\mu^2+c_1\mu+c_2)\mathbf{A}^{n-3}+\ldots+(\mu^{n-1}+c_1\mu^{n-2}+\ldots+c_{n-1})\mathbf{I}\,.$$

Thus, for the planar case $n=2$, the polynomial (3.4) takes the form

$$\Phi(\mu,t)=t\mu^3+(G(X_0)+c_1t)\mu^2$$

$$+\left\{2\,B^T(\mathbf{A}+c_1\mathbf{I})X_0+c_2(t-X_0^TX_0)-c_1+B^TB\right\}\mu+B^T(\mathbf{A}+c_1\mathbf{I})B-c_2$$

while formula (3.3) is transformed into

$$X_*=\frac{1}{f(\mu_*)}\left[(\mathbf{A}+c_1\mathbf{I})+\mu_*\mathbf{I}\right](B+\mu_*X_0)\,.$$

Further simplification of the expressions for distance equation and for the nearest point coordinates can be performed for the case where the ellipse equation is represented in canonical form. For deducing the results of two following corollaries we also apply the equality of Corollary 2.1.

Corollary 3.1. *For the point $X_0=[x_0,y_0]$ and the ellipse*

$$G(x,y):=x^2/a^2+y^2/b^2-1=0 \tag{3.5}$$

the distance equation can be constructed in the form (3.1) where

$$\Phi(\mu,t)=\mu^3+A_1\mu^2+A_2\mu+A_3 \tag{3.6}$$

and

$$\begin{aligned}A_1 &= x_0^2+y_0^2-a^2-b^2-t\,,\\ A_2 &= a^2b^2\left\{\left(\frac{1}{a^2}+\frac{1}{b^2}\right)t-G(x_0,y_0)\right\}\,,\\ A_3 &= -a^2b^2\,t\,.\end{aligned}$$

Using formula (2.3), one can express $\mathcal{F}(t)$ in terms of the coefficients of $\Phi(\mu, t)$, while its further expansion in powers of t is as follows:

$$\mathcal{F}(t) \equiv B_0 t^4 + B_1(x_0, y_0)t^3 + B_2(x_0, y_0)t^2 + B_3(x_0, y_0)t + B_4(x_0, y_0) \quad (3.7)$$

where $B_0 := L^2$,

$$\begin{aligned}
B_1(x_0, y_0) &:= -2\,L\Big\{L(a^2 + b^2 + x_0^2 + y_0^2) + a^2 y_0^2 - b^2 x_0^2\Big\}\,, \\
B_2(x_0, y_0) &:= 6\,L[a^4 y_0^2 + a^2 y_0^4 - b^4 x_0^2 - b^2 x_0^4 + L(a^2 b^2 + x_0^2 y_0^2)] \\
&\quad +[L^2 - (a^2 x_0^2 + b^2 y_0^2)]^2\,, \\
B_3(x_0, y_0) &:= -2\,a^2 b^2 \Big\{a^2 b^2 T_0 G_0^2 - \Big[(a^2 + b^2)T_0^2 + 3\,a^2 b^2 T_0 - 6\,a^4 b^4 S_{4,0}\Big] G_0 \\
&\qquad +2\,a^2 b^2 T_0^2 S_{4,0}\Big\}\,, \qquad (3.8) \\
B_4(x_0, y_0) &:= a^4 b^4 G_0^2 \left(T_0^2 + 4\,a^2 b^2 G_0\right)\,. \qquad (3.9)
\end{aligned}$$

and

$$L := a^2 - b^2,\ G_0 := G(x_0, y_0),\ T_0 := x_0^2 + y_0^2 - a^2 - b^2,\ S_{4,0} := \frac{x_0^2}{a^4} + \frac{y_0^2}{b^4}\,.$$

Once the minimal positive zero t_* of the distance equation is evaluated, one can express the coordinates of the point in the ellipse nearest to $[x_0, y_0]$ by:

$$x_* = \frac{a^2 x_0}{a^2 - \mu_*},\ y_* = \frac{b^2 y_0}{b^2 - \mu_*}\,. \quad (3.10)$$

Here μ_* stands for the multiple zero of a cubic polynomial $\Phi(\mu, t_*)$ defined by (3.6). It can be calculated by formula (2.4) where substitution $t = t_*$ should be made into the expressions for A_1, A_2, A_3 taken from Corollary 3.1.

Remark 1. *In [13] some properties of the distance equation given by (3.7) have been discussed including the structure of its set of real zeros. For some reason to be clarified below, we would like to underline one point of that discussion: the minimal positive zero of the distance equation does not always corresponds to the square of the distance from the point to ellipse. Indeed, the assumption of simplicity of the minimal positive zero of the distance equation stated in Theorem 3.1 is an essential one. For some choices of the point* $[x_0, y_0]$*, the minimal positive zero of the distance equation is generated by a pair of* imaginary *points in the ellipse.*

Corollary 3.2. *For the point $X_0 = [x_0, y_0, z_0]$ and the ellipsoid*

$$G(x, y, z) := x^2/a^2 + y^2/b^2 + z^2/c^2 - 1 = 0 \tag{3.11}$$

the distance equation can be constructed in the form (3.1) where

$$\Phi(\mu, t) = \mu^4 + A_1\mu^3 + A_2\mu^2 + A_3\mu + A_4\,. \tag{3.12}$$

Here

$$\begin{aligned}
A_1 &:= x_0^2 + y_0^2 + z_0^2 - t - a^2 - b^2 - c^2, \\
A_2 &:= a^2b^2c^2\Big\{\left(\frac{1}{b^2c^2} + \frac{1}{a^2c^2} + \frac{1}{a^2b^2}\right)t + \left(\frac{x_0^2}{a^4} + \frac{y_0^2}{b^4} + \frac{z_0^2}{c^4}\right) \\
&\qquad - \left(\frac{1}{a^2} + \frac{1}{b^2} + \frac{1}{c^2}\right)G_0\Big\}, \\
A_3 &:= a^2b^2c^2\Big\{G_0 - \left(\frac{1}{a^2} + \frac{1}{b^2} + \frac{1}{c^2}\right)t\Big\}, \\
A_4 &:= a^2b^2c^2\,t
\end{aligned}$$

and $G_0 := G(x_0, y_0, z_0)$.

Proof. The result can be obtained via application of formula (3.4)

$$\begin{aligned}
\Phi(\mu, t) \equiv & -\left(\mu - \frac{1}{a^2}\right)\left(\mu - \frac{1}{b^2}\right)\left(\mu - \frac{1}{c^2}\right)\left(\mu(t - x_0^2 - y_0^2 - z_0^2) - 1\right) \\
& -\mu^2\Big[x_0^2\left(\mu - \frac{1}{b^2}\right)\left(\mu - \frac{1}{c^2}\right) + y_0^2\left(\mu - \frac{1}{a^2}\right)\left(\mu - \frac{1}{c^2}\right) \\
& \qquad + z_0^2\left(\mu - \frac{1}{a^2}\right)\left(\mu - \frac{1}{b^2}\right)\Big].
\end{aligned}$$

and further reversion the order of the coefficients of powers of μ (Corollary 2.1). □

Using formula (2.5), one can express $\mathcal{F}(t)$ in terms of the coefficients of $\Phi(\mu, t)$, while its further expansion in powers of t becomes rather cumbersome. It is represented here by just a few of its terms (which will be of use in the next section)

$$\mathcal{F}(t) \equiv L^2t^6$$

$$-2\,L\Big\{a^2(b^2-c^2)(2\,a^2-b^2-c^2)(z_0^2+y_0^2)+b^2(c^2-a^2)(2\,b^2-c^2-a^2)(x_0^2+z_0^2)$$
$$+c^2(a^2-b^2)(2\,c^2-b^2-a^2)(x_0^2+y_0^2)+L(a^2+b^2+c^2)\Big\}t^5$$
$$+\dots$$
$$+a^6b^6c^6(k_{1,0}+k_{1,1}G_0+k_{1,2}G_0^2+k_{1,3}G_0^3+k_{1,4}G_0^4)t$$
$$+a^6b^6c^6(k_{0,0}+k_{0,1}G_0+k_{0,2}G_0^2+k_{0,3}G_0^3)G_0^2\,. \tag{3.13}$$

Here

$$L:=(a^2-b^2)(a^2-c^2)(b^2-c^2)\,,$$

$$k_{1,0} := -4\,a^2b^2c^2S_{4,0}^3(T_0^2-4\,a^2b^2c^2S_{4,0})\,, \tag{3.14}$$
$$k_{1,1} := 2\,S_{4,0}\{9\,T_0^3+5(a^2b^2+a^2c^2+b^2c^2)T_0^2S_{4,0} -28a^2b^2c^2(a^2b^2+a^2c^2+b^2c^2)S_{4,0}^2-40\,a^2b^2c^2T_0S_{4,0}\}\,, \tag{3.15}$$
$$k_{0,0} := a^2b^2c^2S_{4,0}^2(T_0^2-4\,a^2b^2c^2S_{4,0})\,, \tag{3.16}$$
$$k_{0,1} := -4\,T_0^3-2\,(a^2b^2+a^2c^2+b^2c^2)T_0^2S_{4,0} + 12\,(a^2b^2+a^2c^2+b^2c^2)a^2b^2c^2S_{4,0}^2+18\,a^2b^2c^2T_0S_{4,0} \tag{3.17}$$

and $T_0:=x_0^2+y_0^2+z_0^2-a^2-b^2-c^2$,

$$S_{4,0}:=x_0^2/a^4+y_0^2/b^4+z_0^2/c^4\,. \tag{3.18}$$

The non-identified coefficients $\{k_{ij}\}$ in (3.13) are just polynomials in x_0,y_0,z_0.

Once the minimal positive zero t_* of the distance equation is evaluated, the coordinates of the point in the ellipsoid nearest to $[x_0,y_0,z_0]$ can be expressed from (3.3) as:

$$x_*=\frac{a^2x_0}{a^2-\mu_*},\ y_*=\frac{b^2y_0}{b^2-\mu_*},\ z_*=\frac{c^2z_0}{c^2-\mu_*}\,. \tag{3.19}$$

Here μ_* stands for the multiple zero of a quartic polynomial $\Phi(\mu,t_*)$ defined by (3.12). This zero can be calculated by formula (2.6) in which the values for A_1, A_2, A_3, A_4 are taken from Corollary 3.2 with substitution $t=t_*$ made into their expressions.

Remark 2. *The results of the present section can be extended to the problem of finding* the farthest *distance from X_0 to the points in the ellipse or ellipsoid. For this aim, in the previous results of the present section, one should take t_* to be* the greatest *positive zero for the corresponding distance equation.*

Example 3.1. *For the point* $[6,-2,5]$*, find the nearest and the farthest point in the ellipsoid* $x^2/4+y^2/16+z^2/49=1$.

Solution. Compute the distance equation via Corollary 3.2:

$$19847025\,t^6-8393808060\,t^5+1317736785456\,t^4-103262746605120\,t^3$$

$$+4327358033988864\,t^2-91883501048862720\,t+757148717189025792=0\ .$$

It has exactly two positive zeros, namely $t_1\approx 21.63634$, $t_2\approx 186.72961$. Therefore the distance from the given point to the ellipsoid equals $\sqrt{t_1}\approx 4.65149$ while the distance to the farthest point in ellipsoid equals $\sqrt{t_2}\approx 13.66490$. Compute the multiple zero for polynomial $\Phi(\mu,t_j)$ by formula (2.6):

$$\mu_j=p(u_j)/q(u_j)$$

where

$$\begin{aligned} p(u) &:= 2\big(94581\,u^5-26525988\,u^4+2036643696\,u^3\\ &\quad -41839985024\,u^2-742245672192\,u+16044295020544\big)\,,\\ q(u) &:= 46683\,u^5-16504668\,u^4+1934686224\,u^3\\ &\quad -101284594048\,u^2+2376476312064\,u-19462452484096\,. \end{aligned}$$

This yields: $\mu_1\approx -11.70096$, $\mu_2\approx 84.64247$. Formulas (3.19) give one the coordinates of the nearest $X_1\approx[1.52857,-1.15519,4.03618]$ and the farthest $X_2\approx[-0.29761,0.46618,-6.87382]$ point in the considered ellipsoid. □

4. Distance Equation Zero Approximations

In the previous sections we have presented an analytical solution for the distance evaluation problem. The stated problem is reduced to that of solving the distance equation. For any specialization of parameters (i.e., the given point coordinates and the coefficients of the ellipse or ellipsoid equations), this can be done numerically. However, to resolve these equations analytically and to obtain an explicit expression for the distance as a function of the parameters is not a trivial task. This is possible for the point-to-ellipse problem since the corresponding distance equation (3.7) is generically (i.e. if an ellipse is not a circle) of the degree 4 and therefore it can be resolved by radicals[1]. As for the case of the point-to-ellipsoid distance equation represented by (3.13), its degree generically (i.e. if an ellipsoid is not a rotational one) equals to 6, and therefore one cannot expect its solubility by radicals.

One of the possible approaches for evaluation of a zero of an algebraic equation

$$h_0 t^m + h_1 t^{m-1} + \ldots + h_{m-1} t + h_m = 0$$

consists in its linearization in a vicinity of a zero being searched. For instance, would this zero t_* be close to 0, one may expect that the approximation

$$t_* \approx -h_m / h_{m-1}$$

is a satisfactory one. Being applied for the distance equation, this reason leads one to analytical distance approximation formulas. For the planar case, such a formula

$$d_{HO} := \sqrt{-B_4/B_3}\,, \tag{4.1}$$

is suggested in [14]; here B_3 and B_4 are defined by (3.8) and (3.9) respectively. Note that this formula fails when the radicand happens to be negative.

Example 4.1. *For the ellipse $x^2/a^2 + y^2/25 = 1$ formula (4.1) provides the following approximations of the distance to it from the point $[a + 0.5, 0]$:*

a	5	10	30	60	67	68	69
$d_{HO} \approx$	0.499	0.511	0.605	1.245	2.786	4.282	n/a

For $a = 69$ the radicand is negative. Even replacement of the latter by its absolute value does not lead to a satisfactory result.

Explanation for this fault is as follows. Distance equation (3.7) possesses two positive zeros and one double zero

$$z_1 = \frac{1}{4},\ z_2 = \left(2\,a + \frac{1}{2}\right)^2,\ z_3 = z_4 = -\frac{25}{4} \cdot \frac{4\,a + 101}{a^2 - 25}$$

which is negative when $a > 5$. Vieta's formulas provide the equality

$$d_{HO}^2 = \frac{1}{1/z_1 + 1/z_2 + 1/z_3 + 1/z_4}$$

[1]At least, in principle, i.e. ignoring the obstacles of the *casus irreducibilis* when the real zero is represented as a combination of radicals of imaginary radicands [16].

which, in the case where z_1 is much smaller than the absolute values of other zeros, yields a satisfactory approximation. However, for $a = 69$ one has $1/z_1 + 1/z_3 + 1/z_4 \approx -0.01994$, and thus d^2_{HO} fails to approximate z_1. □

This example is an illustration of the degenerate case stated in the assumption of Theorem 3.1, namely that one where the distance equation possesses a multiple real zero. As it was declared in Remark 1, it might happened that this zero is generated by a pair of *imaginary* points in the considered ellipse or ellipsoid. It is exactly the case of example (4.1).

Theorem 4.1. *The radicand in (4.1) is negative iff the coefficient B_3 given by (3.8) is positive.*

Proof Follows from an alternative representation for the coefficient B_4 defined by (3.9):

$$B_4 \equiv a^4 b^4 G_0^2 \left[(x_0^2 - y_0^2 + b^2 - a^2)^2 + 4\, x_0^2 y_0^2 \right] .$$

Therefore, it is nonnegative for any specialization of parameters. □

From this theorem it follows that applicability of formula (4.1) depends crucially on the relative position of the curve

$$B_3(x, y) = 0 \tag{4.2}$$

defined by (3.8) w.r.t. the ellipse (3.5). Depending on the parameter specializations, this 6th order curve contains from 2 to 4 ovals (branches).

Theorem 4.2. *For any specialization of parameters a and b with $a > b$, the curve (4.2) has two ovals lying inside the ellipse. For $b < a \le \sqrt{2}b$, the curve does not have any other oval. For $a > \sqrt{2}b$, the curve possesses two extra unclosed ovals lying outside the ellipse.*

Proof. It can be established that any possible oval of the curve (4.2) has an intersection point with the x-axis. These points $[\widehat{x}_j, 0]$ can be determined from the equation

$$B_3(x, 0) \equiv 2\, b^2 (x^2 - a^2 + b^2)((a^2 - 2\, b^2) x^4 + (-b^4 + 2\, a^2 b^2 - 2\, a^4) x^2 + a^6 - a^2 b^4) = 0 .$$

One has

$$\widehat{x}_1^2 \ := \ a^2 - b^2 \ , \tag{4.3}$$

$$\widehat{x}_2^2 \ := \ \frac{a^2}{2} + \frac{a^4 + b^4 - b^2\sqrt{12\, a^4 - 12\, a^2 b^2 + b^4}}{2(a^2 - 2\, b^2)} \ , \tag{4.4}$$

$$\widehat{x}_3^2 \ := \ \frac{a^2}{2} + \frac{a^4 + b^4 + b^2\sqrt{12\, a^4 - 12\, a^2 b^2 + b^4}}{2(a^2 - 2\, b^2)} \ , \tag{4.5}$$

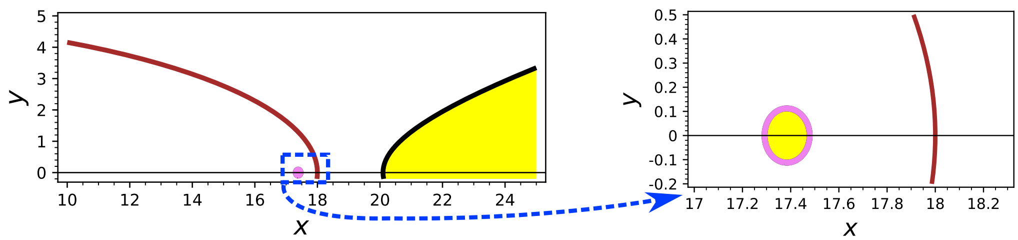


Figure 4.1: Domains of failure of the approximation (4.1) for the ellipse $x^2/324+y^2/25=1$ (thin red line).

and it can be easily verified that the right-hand side of (4.4) is positive for any values of parameters a and b such that $a^2 \neq 2\,b^2$ while the right-hand side of (4.5) is negative (positive) iff $a^2-2\,b^2$ is negative (positive). If one takes for $\widehat{x}_j$ the positive values defined by (4.3)–(4.5), then $0<\widehat{x}_1<\widehat{x}_2<a<\widehat{x}_3$. The points $[\widehat{x}_1,0]$ and $[\widehat{x}_2,0]$ are the points in the closed oval of the curve (4.2) internal to the ellipse. □

Example 4.2. *For the ellipse $x^2/324+y^2/25=1$, the curve (4.2) is displayed*[2] *in Fig. 4.1.*

The points $X_0=[x_0,y_0]$ close to the curve (4.2) break the formula (4.1) confidence. For instance, the points close to a tiny oval internal to the ellipse yield

$[x_0,y_0]$	$[17.20,0]$	$[17.30,0]$	$[17.40,0]$	$[17.48,0]$	$[17.50,0]$	$[17.60,0]$
d_{HO}	0.33	n/a	n/a	4.60	1.19	0.50

with the n/a meaning the negative sign of the involved radicand.

On the contrary, branches of the curve (4.2) lying outside the ellipse border the infinite domains of failure (inapplicability) of the approximation (4.1). Approximation error grows drastically when $X_0=[x_0,y_0]$ falls close to these branches: $d_{HO}(20.00,0.00)\approx 8.49, d_{HO}(20.10,0.10)\approx 50.62$ etc. □

In Section 5 we will give a quantitative error analysis for the approximation (4.1). It is evident that this task correlates somehow with the problem

[2]Since all the plots in the paper possess a symmetry property w.r.t. the axes, they are displayed only in the first quadrant of the Cartesian plane.

of evaluation of the qualitative behavior of the locus of the curve (4.2) in its dependency on the parameters a and b. We sketch briefly the dynamics of this locus for a fixed a, and b decreasing from a to 0 (i.e. with the increasing of the ellipse eccentricity). The closed tiny oval lying inside the ellipse appears from the origin and moves along the x-axis toward $[a, 0]$ with its size increasing until b reaches the value $a/\sqrt{2}$. On the passage of this bifurcation value, a new unclosed oval of the curve is generated at the infinite point $[+\infty, 0]$. For further decreasing value of b, this oval moves along the x-axis toward $[a, 0]$ with its nearest point to the ellipse at $[\widehat{x}_3, 0]$. As the ratio $\tau := b/a$ tends to 0, the asymptotical behavior of the positive values $\{\widehat{x}_j\}_{j=1}^3$ defined by (4.3)–(4.5) is as follows

$$\widehat{x}_1 = a\left(1 - \frac{1}{2}\tau^2\right), \ \widehat{x}_2 \sim a\left(1 + \frac{1-\sqrt{3}}{2}\tau^2\right), \ \widehat{x}_3 \sim a\left(1 + \frac{1+\sqrt{3}}{2}\tau^2\right).$$

For the parameter values such that $b > a/\sqrt{2}$, the internal oval can be approximated satisfactory by the circle centered at $[(\widehat{x}_2 + \widehat{x}_1)/2, 0]$ and with the diameter $(1 - \sqrt{3}/2)b^2/a$. Thus, it shrinks as b tends to zero. For any parameter values, this *stain of outliers* is small enough compared with the size of the ellipse and therefore can be neglected without aftereffects for the applications like the ellipse fitting problem mentioned in Introduction. On the contrary, this is not the case when dealing with the external to the ellipse branch of the curve (4.2) since its destructive effect onto the approximation (4.1) increases when it approaches the ellipse.

We now turn to an alternative approach for finding approximation of a zero of an algebraic equation in terms of its coefficients. It consists in finding an expansion of the zero in an appropriate power series. For the case of distance equation, let us choose the latter to be the one in powers of the algebraic distance, namely the value $G(X_0)$. This value, in a neighborhood of the considered ellipse or ellipsoid, can be treated as a small parameter. We restrict our treatment with computing the first three terms of this expansion:

$$t_* = \ell_1 G(X_0) + \ell_2 G^2(X_0) + \ell_3 G^3(X_0) + \dots \tag{4.6}$$

To determine its coefficients ℓ_1, ℓ_2 and ℓ_3, substitute this expansion into the distance equation (3.1), expand the result in powers of $G(X_0)$ and equate then the coefficients of $G(X_0), G^2(X_0), G^3(X_0)$ to zero. Our successes in completing these plan are restricted to the cases of $\mathbb{R}^2$ and $\mathbb{R}^3$.

Theorem 4.3. *For the ellipsoid in* $\mathbb{R}^3$ *represented in canonical form (3.11), the first two approximations for the distance to it from the point* $X_0 = [x_0, y_0, z_0]$ *are given as*

$$d_{(1)} = \frac{1}{2}\frac{|G(X_0)|}{\sqrt{S_{4,0}}} \tag{4.7}$$

and

$$d_{(2)} = d_{(1)}\sqrt{1 + \frac{1}{2}\frac{S_{6,0}}{S_{4,0}^2}G(X_0)} \tag{4.8}$$

provided the radicand is nonnegative. Here

$$S_{4,0} \stackrel{(3.18)}{=} x_0^2/a^4 + y_0^2/b^4 + z_0^2/c^4;\ S_{6,0} := x_0^2/a^6 + y_0^2/b^6 + z_0^2/c^6\,.$$

The counterparts of these formulas for the case of planar ellipse (3.5) and the point $X_0 = [x_0, y_0]$ *are derived by setting* $z_0 = 0$.

Proof. We restrict ourselves with the case of $\mathbb{R}^3$. Substitute (4.6) into (3.7) and equate the coefficients of powers of $G_0 = G(X_0)$ to zero. One gets $\ell_1 = 0$ and therefore the coefficients of powers of $t^2, t^3, \dots, t^6$ in (3.7) do not influence the values for ℓ_2 and ℓ_3. The latter ones can be determined from the following linear relations

$$k_{1,0}\ell_2 + k_{0,0} = 0,\ k_{1,0}\ell_3 + k_{1,1}\ell_2 + k_{0,1} = 0$$

with the involved coefficients $\{k_{ij}\}$ defined by (3.14)–(3.17). Resolving these equations results in

$$\ell_2 = -\frac{k_{0,0}}{k_{1,0}} = \frac{1}{4\,S_{4,0}},\ \ell_3 = \ell_2\left(\frac{k_{0,1}}{k_{0,0}} - \frac{k_{1,1}}{k_{1,0}}\right)$$

wherefrom immediately follows (4.7) for the first approximation of the $\sqrt{t_*}$. The expression for ℓ_3 needs some extra efforts:

$$\frac{k_{0,1}}{k_{0,0}} - \frac{k_{1,1}}{k_{1,0}} = \frac{1}{2}\cdot\frac{(a^2b^2 + a^2c^2 + b^2c^2)S_{4,0} + T_0}{a^2b^2c^2S_{4,0}^2}$$

and it turns out that the numerator of the last fraction can be reduced by an additional term containing G_0 since

$$(a^2b^2 + a^2c^2 + b^2c^2)S_{4,0} + T_0 \equiv (a^2 + b^2 + c^2)G_0 + a^2b^2c^2S_{6,0}\,.$$

This completes the proof of (4.8). □

Example 4.3. *Find distance approximations for the problem of Example 3.1.*

Solution. One has:

$$G(X_0) \approx 8.76020,\ d_{(1)} \approx 2.90332,\ d_{(2)} \approx 3.52799,\ d_{HO} \approx 2.87059\,;$$

thus $d_{(2)}$ approximates the distance $d \approx 4.651488$ within 25% error.

Let us test some extra points lying at the distance $d \approx 1$ from the considered ellipsoid,

X_0	$\begin{bmatrix} 1 \\ 1 \\ 7.49 \end{bmatrix}$	$\begin{bmatrix} 0.5 \\ 2 \\ 7.26 \end{bmatrix}$	$\begin{bmatrix} 0.5 \\ 2 \\ 2.53 \end{bmatrix}$	$\begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \\ 8 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \\ 6 \end{bmatrix}$
$G(X_0)$	0.46	0.39	-0.55	1.25	-0.75	0.31	-0.26
d_{HO}	1.10	2.46	0.78	0.93	0.89	n/a	0.68
$d_{(1)}$	0.58	0.71	1.51	0.83	1.5	0.94	1.08
$d_{(2)}$	0.85	0.97	n/a	0.94	n/a	0.99	0.98

here n/a stands for a negative value of the radicand. □

Once the two approximations for the distance are obtained for quadric given in canonical form, one can extend these results to the case of a quadric represented in its general form (1.4).

Theorem 4.4. *For the distance from the point $X_0 \neq -\mathbf{A}^{-1}B$ to the quadric (1.4) in $\mathbb{R}^2$ and $\mathbb{R}^3$ the first approximation formula is given by*

$$d_{(1)} = \frac{1}{2} \cdot \frac{|G(X_0)|}{\sqrt{(\mathbf{A}X_0 + B)^T(\mathbf{A}X_0 + B)}}\,, \tag{4.9}$$

while the second one by

$$d_{(2)} = d_{(1)}\sqrt{1 + \frac{1}{2} \cdot \frac{(\mathbf{A}X_0 + B)^T\mathbf{A}(\mathbf{A}X_0 + B)}{[(\mathbf{A}X_0 + B)^T(\mathbf{A}X_0 + B)]^2}G(X_0)}\,. \tag{4.10}$$

Proof. We restrict ourselves here with the proof of both formulas for the case $\mathbb{R}^3$. Let us start with the case when $B = \mathbb{O}$; thus matrix $\mathbf{A}$ is assumed to be positive definite. The distance from the point X_0 to the ellipsoid $X^T\mathbf{A}X = 1$ is unaltered under the transformation $Y = QX$ with an orthogonal matrix

Q: it equals to the distance from $Y_0 = QX_0$ to $Y^TQ\mathbf{A}Q^TY = 1$. Choose this transformation with the aim to reduce the ellipsoid to canonical form:

$$Q\mathbf{A}Q^T = \mathbf{A}_{diag} = \left(\begin{array}{ccc} 1/a^2 & 0 & 0 \\ 0 & 1/b^2 & 0 \\ 0 & 0 & 1/c^2 \end{array} \right) .$$

Here $1/a^2, 1/b^2, 1/c^2$ stand for the eigenvalues of the matrix $\mathbf{A}$. Starting with this matrix form, one can transform the representation (4.7):

$$d_{(1)} = \frac{|Y_0^T \mathbf{A}_{diag} Y_0 - 1|}{2\sqrt{Y_0^T \mathbf{A}_{diag}^2 Y_0}}$$

$$= \frac{|Y_0^T Q\mathbf{A}Q^T Y_0 - 1|}{2\sqrt{Y_0^T (Q\mathbf{A}Q^T)^2 Y_0}} = \frac{|Y_0^T Q\mathbf{A}Q^T Y_0 - 1|}{2\sqrt{Y_0^T Q\mathbf{A}^2 Q^T Y_0}} = \frac{|X_0^T \mathbf{A} X_0 - 1|}{2\sqrt{X_0^T \mathbf{A}^2 X_0}} . \qquad (4.11)$$

Similarly, one can find the counterpart of the approximation (4.8) for the ellipsoid $X^T\mathbf{A}X = 1$:

$$d_{(2)} = d_{(1)} \sqrt{1 + \frac{1}{2} \cdot \frac{X_0^T \mathbf{A}^3 X_0}{X_0^T \mathbf{A}^2 X_0} (X_0^T \mathbf{A} X_0 - 1)} . \qquad (4.12)$$

The case where $B \neq \mathbb{O}$ can be reduced to the previous one with the aid of transformation

$$X = Y + X_c, \text{ where } X_c := -\mathbf{A}^{-1}B \qquad (4.13)$$

denotes the ellipsoid center. The ellipsoid equation takes now the form

$$Y^T\widetilde{\mathbf{A}}Y = 1 \text{ with } \widetilde{\mathbf{A}} := -\frac{\mathbf{A}}{G(X_c)} = \frac{\mathbf{A}}{B^T\mathbf{A}^{-1}B + 1} .$$

Distance to this ellipsoid from the point $Y_0 = X_0 - X_c$ coincides with that from X_0 to the ellipsoid (1.4). One has:

$$\widetilde{\mathbf{A}}Y_0 = -\frac{1}{G(X_c)}\mathbf{A}(X_0 - X_c) = -\frac{1}{G(X_c)}(\mathbf{A}X_0 + B) , \qquad (4.14)$$

$$\widetilde{\mathbf{A}}^2 Y_0 = \frac{1}{G^2(X_c)}\mathbf{A}^2(X_0 - X_c) = \frac{1}{G^2(X_c)}\mathbf{A}(\mathbf{A}X_0 + B) . \qquad (4.15)$$

With the aid of these relations, transform the expressions for $Y_0^T \widetilde{\mathbf{A}} Y_0, Y_0^T \widetilde{\mathbf{A}}^2 Y_0$ and $Y_0^T \widetilde{\mathbf{A}}^3 Y_0$. From

$$Y_0^T \widetilde{\mathbf{A}} Y_0 \stackrel{(4.14)}{=} -\frac{1}{G(X_c)}(X_0 - X_c)^T(\mathbf{A}X_0 + B)$$

$$= -\frac{1}{G(X_c)}(X_0^T \mathbf{A} X_0 + X_0^T B - X_c^T \mathbf{A} X_0 - X_c^T B)$$

$$\stackrel{(4.13)}{=} -\frac{1}{G(X_c)}(X_0^T \mathbf{A} X_0 + 2B^T X_0 + B^T \mathbf{A}^{-1} B)$$

$$= -\frac{1}{G(X_c)}(G(X_0) + B^T \mathbf{A}^{-1} B + 1) = -\frac{G(X_0)}{G(X_c)} + 1\,;$$

and

$$Y_0^T \widetilde{\mathbf{A}}^2 Y_0 \stackrel{(4.14)}{=} \frac{1}{G^2(X_c)}(X_0 - X_c)^T \mathbf{A}(\mathbf{A}X_0 + B)$$

$$= \frac{1}{G^2(X_c)}(X_0^T \mathbf{A} - X_c^T \mathbf{A})(\mathbf{A}X_0 + B) \stackrel{(4.13)}{=} \frac{1}{G^2(X_c)}(\mathbf{A}X_0 + B)^T(\mathbf{A}X_0 + B)$$

follows formula (4.11). Transformation of

$$Y_0^T \widetilde{\mathbf{A}}^3 Y_0 = Y_0^T \widetilde{\mathbf{A}} \cdot \widetilde{\mathbf{A}}^2 Y_0 \stackrel{(4.14),(4.15)}{=} -\frac{1}{G^3(X_c)}(\mathbf{A}X_0 + B)^T(\mathbf{A}^2 X_0 + \mathbf{A}B)$$

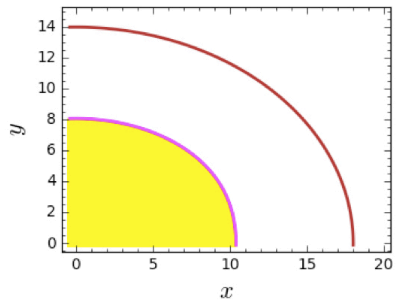
$$= -\frac{1}{G^3(X_c)}(\mathbf{A}X_0 + B)^T \mathbf{A}(\mathbf{A}X_0 + B)$$

completes the proof of (4.12). □

The geometric interpretation for the approximation (4.9) is as follows. This value coincides with the distance from the point X_0 to the linear manifold obtained by linearization of $G(X)$ at the point X_0:

$$G(X_0) + \left.\frac{\mathbf{D}\,G}{\mathbf{D}\,X}\right|_{X=X_0}(X - X_0) = 0\ .$$

Here the row vector $\mathbf{D}\,G/\mathbf{D}\,X|_{X=X_0}$ stands for the gradient of the function $G(X)$ calculated in the point X_0. This formula was suggested in [5] for the distance approximation from a point X_0 to arbitrary algebraic manifold $G(X) = 0$ in $\mathbb{R}^n$. However in subsequent papers, accuracy of the estimation has been evaluated only empirically since the tight bounds for it (the worst possible cases) need the exact analytical formula for the distance function.



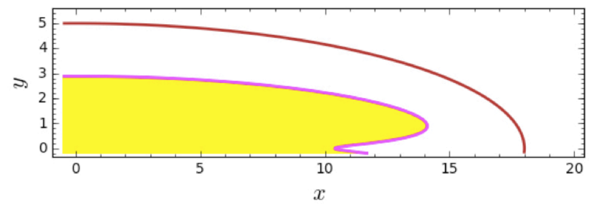


Figure 4.2: Domain of failure for the approximation (4.10) for the ellipse $x^2/324+y^2/b^2 = 1$ (thin red line): $b = 14$ (left) and $b = 5$ (right).

In the next section we will be focused onto resolving this problem for quadrics in $\mathbb{R}^2$ and in $\mathbb{R}^3$ while to conclude the present one we have to establish the domain of applicability of the approximation (4.10). The radicand in its right-hand side might be negative. We restrict here our consideration with the planar case and the ellipse equation in canonical form (3.5). The structure of the 4th order algebraic curve

$$H(x,y) := 2\left(\frac{x^2}{a^4}+\frac{y^2}{b^4}\right)^2 + \left(\frac{x^2}{a^6}+\frac{y^2}{b^6}\right)\left(\frac{x^2}{a^2}+\frac{y^2}{b^2}-1\right) = 0 \qquad (4.16)$$

is simpler that its counterpart (4.2). It consists of a single oval lying inside the ellipse. For $a = b$ this is just a circle $x^2 + y^2 = b^2/3$. When b decreases while a remains fixed, the curve flattens along the x-axis and intersects the axes at $[0, \pm b/\sqrt{3}]$ and $[\pm a/\sqrt{3}, 0]$ (Fig. 4.2). The points $X_0 = [x_0, y_0]$ in the curve maximize the error for the approximation (4.10). However, it should be emphasized that this error is not an infinite one (in comparison with the singularity of the approximation (4.1)).

To summarize the domains of failure for the point-to-ellipse distance approximation formulas presented above, we compile the following table. Here the ellipse is represented in canonical form (3.5) and the point does not coincide with the origin (where all the formulas fail).

approximation formula	radicand singularity	radicand negativity	location w.r.t. ellipse
(4.1) (Harker & O'Leary)	yes $B_3(x,y) \overset{(3.8)}{=} 0$	yes $B_3(x,y) \overset{(3.8)}{>} 0$	both in and outside
(4.7) (Sampson)	no	no	n/a
(4.8)	no	yes, $H(x,y) \overset{(4.16)}{<} 0$	inside

5. Error Estimation for Approximations

To estimate the error of the distance approximation by formulas (4.9) or (4.10), let us specialize the values for these approximations and evaluate the maximal and minimal distances for the points in the obtained manifolds from the approximated quadric (1.4). We will call by the *maximal* (or the *minimal*) *deviation* of the manifold from the quadric (1.4) the maximal (or, respectively, the minimal) distance between all the pairs of the nearest points in these manifolds.

We restrict ourselves here to the case of the quadric represented in canonical form. Consider first Sampson's approximation (4.7). For the planar case the level curves of $d^2_{(1)}(x,y) = h$ can be represented as the 4th order algebraic curves:

$$K_{\sqrt{h}}(x,y) := \left(\frac{x^2}{a^2} + \frac{y^2}{b^2} - 1\right)^2 - 4\,h\left(\frac{x^2}{a^4} + \frac{y^2}{b^4}\right) = 0\,. \tag{5.1}$$

Here the positive parameter $h > 0$ has the meaning of the squared "approximate distance".

Theorem 5.1. *Let $a > b > \sqrt{h}$. The values of the minimal and the maximal deviation of (5.1) from the ellipse (3.5) are either among the numbers*

$$\left\{\sqrt{h} \pm \left(a - \sqrt{a^2+h}\right)\ ;\ \sqrt{h} \pm \left(b - \sqrt{b^2+h}\right)\right\} \tag{5.2}$$

or among the square roots of positive zeros of the equation

$$\mathcal{K}_h(t) := C_0 t^4 + C_1 t^3 + C_2 t^2 + C_3 t + C_4 = 0\,. \tag{5.3}$$

Here

$$
\begin{aligned}
C_0 &= a^4b^4(a^2-b^2)^2,\\
C_1 &= -2\,a^2b^2(a^8+b^8-4\,a^4b^4)h-2\,a^4b^4(a^2+b^2)(2\,a^4-5\,a^2b^2+2\,b^4),\\
C_2 &= (a^8-10\,a^4b^4+b^8)(a^2+b^2)^2h^2+6\,a^4b^4(a^2+b^2)(a^4-7\,a^2b^2+b^4)h\\
&\quad -27\,a^8b^8,\\
C_3 &= 2\,a^2b^2h\left\{2\,(a^2+b^2)h+3\,a^2b^2\right\}\left\{(a^2+b^2)^3h+9\,a^4b^4\right\},\\
C_4 &= -a^4b^4h^2\left\{4\,(a^2+b^2)^3h+27\,a^4b^4\right\}.
\end{aligned}
$$

Proof. We set the problem of constrained optimization: find the critical values (including the maximal and the minimal ones) of the function

$$(x-\tilde{x})^2+(y-\tilde{y})^2 \quad \text{subject to} \quad K_{\sqrt{h}}(x,y)=0,\ G(\tilde{x},\tilde{y})=0\ .$$

For its solution, we utilize the Lagrange multipliers method. We first construct the Lagrange function

$$(x-\tilde{x})^2+(y-\tilde{y})^2-\lambda_1 K_{\sqrt{h}}(x,y)-\lambda_2 G(\tilde{x},\tilde{y})$$

and then equate to zero its derivatives w.r.t. the variables $x, y, \tilde{x}, \tilde{y}, \lambda_1, \lambda_2$:

$$x-\tilde{x} = \frac{2\,\lambda_1 x}{a^2}\left(G(x,y)-\frac{2h}{a^2}\right), \tag{5.4}$$

$$y-\tilde{y} = \frac{2\,\lambda_1 y}{b^2}\left(G(x,y)-\frac{2h}{b^2}\right), \tag{5.5}$$

$$x-\tilde{x} = -\lambda_2\tilde{x}/a^2, \tag{5.6}$$

$$y-\tilde{y} = -\lambda_2\tilde{y}/b^2, \tag{5.7}$$

$$K_{\sqrt{h}}(x,y) = 0, \tag{5.8}$$

$$G(\tilde{x},\tilde{y}) = 0\,. \tag{5.9}$$

Complement the obtained system with the equation

$$t=(x-\tilde{x})^2+(y-\tilde{y})^2 \tag{5.10}$$

which introduces a new variable responsible for the critical values of the distance function. Our aim is to eliminate all the variables from the system (5.4) – (5.10) except for t. To do this, we express $\tilde{x}$ and $\tilde{y}$ from (5.6) and (5.7); the result is similar to (3.10):

$$\tilde{x}=\frac{a^2x}{a^2-\lambda_2}, \quad \tilde{y}=\frac{b^2y}{b^2-\lambda_2}\,. \tag{5.11}$$

Substitute these values into (5.9) and (5.10):

$$\frac{a^2x^2}{(a^2-\lambda_2)^2}+\frac{b^2y^2}{(b^2-\lambda_2)^2}-1 \;=\; 0, \tag{5.12}$$

$$\frac{\lambda_2^2x^2}{(a^2-\lambda_2)^2}+\frac{\lambda_2^2y^2}{(b^2-\lambda_2)^2}-t \;=\; 0\,. \tag{5.13}$$

Substitute next (5.11) into (5.4) and (5.5); the resulting equations can be split into the alternatives:

$$x=0 \quad \text{or} \quad \frac{\lambda_2}{\lambda_2-a^2}=\frac{2\lambda_1}{a^2}\left(G(x,y)-\frac{2\,h}{a^2}\right);$$

$$y=0 \quad \text{or} \quad \frac{\lambda_2}{\lambda_2-b^2}=\frac{2\lambda_1}{b^2}\left(G(x,y)-\frac{2\,h}{b^2}\right).$$

The first parts of these alternatives correspond to the values (5.2). From the second parts, it is possible to eliminate the parameter λ_1:

$$G(x,y)+2\,h(1/a^2+1/b^2-1/\lambda_2)=0\,. \tag{5.14}$$

Equations (5.12), (5.13) and (5.14) compose a linear system w.r.t. x^2 and y^2. The necessary condition for its solubility is that of vanishing the determinant

$$\left|\begin{array}{ccc} a^2/(a^2-\lambda_2)^2 & b^2/(b^2-\lambda_2)^2 & -1 \\ \lambda_2^2/(a^2-\lambda_2)^2 & \lambda_2^2/(b^2-\lambda_2)^2 & -t \\ 1/a^2 & 1/b^2 & -1 \end{array}\right|.$$

This yields an equation

$$F_1(\lambda_2,t):=\lambda_2^3+(a^2+b^2)(2\,h-t)\lambda_2+2\,a^2b^2(t-h)=0\,.$$

Resolve equations (5.12) and (5.13) w.r.t. x^2 and y^2:

$$x^2=\frac{(a^2-\lambda_2)^2(\lambda_2^2-b^2t)}{\lambda_2^2(a^2-b^2)},\; y^2=-\frac{(b^2-\lambda_2)^2(\lambda_2^2-a^2t)}{\lambda_2^2(a^2-b^2)} \tag{5.15}$$

and substitute the result into (5.8):

$$F_2(\lambda_2,t):=\lambda_2^6+2\,a^2b^2t(a^2+b^2)(2\,h-t)\lambda_2^4-4\,a^2b^2(2\,h-t)\lambda_2^3$$

$$+t\left\{(4\,h-t)(a^2+b^2)^2-4\,ha^2b^2\right\}\lambda_2^2+4\,a^2b^2t(a^2+b^2)(2\,h-t)\lambda_2$$

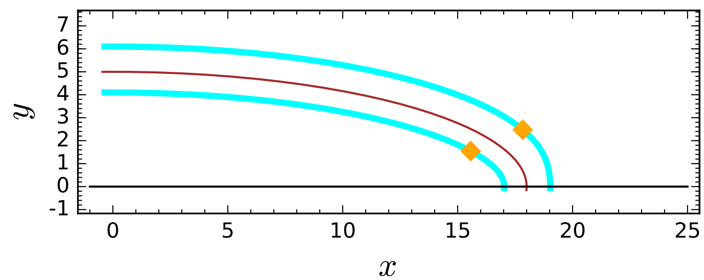


Figure 5.1: Curve $K_1(x, y) = 0$ (cyan) in a vicinity of the ellipse $x^2/324 + y^2/25 = 1$ (thin red line).

$$-4\,a^4b^4t(h-t) = 0\,.$$

The last step is to eliminate the parameter λ_2 from the system of equations $F_1(\lambda_2, t) = 0, F_2(\lambda_2, t) = 0$. We will first simplify one of the involved equations:

$$F_2(\lambda_2, t) - F_1(\lambda_2, t)(F_1(\lambda_2, t) + 2a^2b^2h) \equiv 2\,h\lambda_2\widetilde{F}_2(\lambda_2, t)$$

where

$$\widetilde{F}_2(\lambda_2, t) := -3\,a^2b^2\lambda_2^2 - 2[a^2b^2(2h-t) + h(a^4+b^4)]\lambda_2 + a^2b^2(a^2+b^2)(2\,h-t)\,.$$

Elimination of λ_2 from the system $F_1(\lambda_2, t) = 0, \widetilde{F}_2(\lambda_2, t) = 0$ with the aid of the resultant $\mathcal{R}_{\lambda_2}(F_1, \widetilde{F}_2)$ computation results (up to a factor $4\,a^2b^2$) in the equation (5.3).

Once the positive zero $t = t_*$ of (5.3) is found, the corresponding value for the parameter λ_2 can be established on resolving the equation $\widetilde{F}_2(\lambda_2, t_*) = 0$. Substituting $t_*, \lambda_{2,*}$ into (5.15) for determining a point in the curve (5.1) where the critical value of the distance function is attained. Taking one of these point (for instance, with positive coordinates), one can find the corresponding point in the ellipse (3.5) via formulas (5.11). However, this algorithm fails when $b^2t_* < \lambda_{2,*}^2 < a^2t_*$ since, for such a combination of values t and λ_2, the right-hand sides of (5.15) is negative. □

Example 5.1. *Find the maximal and the minimal deviations of the curves* $\{K_j(x, y) = 0\}_{j=1}^3$ *from the ellipse* $x^2/324 + y^2/25 = 1$.

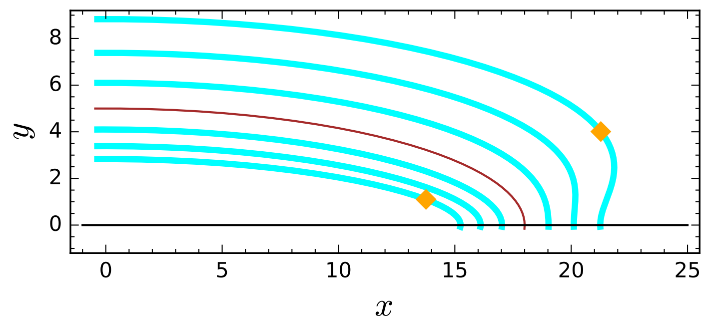


Figure 5.2: Curves $\{K_j(x,y)=0\}_{j=1}^{3}$.

Solution. For $h=1$ equation (5.3) is as follows:

$$5865599610000\,t^4 - 7991709399016200\,t^3 - 108245392226126999\,t^2$$

$$+256343903192012400\,t - 127382090299560000 = 0$$

It possesses the following positive zeros:

$$t_1 \approx 0.74241,\ t_2 \approx 1.37134,\ t_3 \approx 1375.86082\ .$$

The geometry of the curve $K_1(x,y)=0$ is presented in Fig.5.1. It has two branches. The one external to the ellipse has deviations for its points within the interval $(5\sqrt{13}-17, \sqrt{t_2}) \approx (1.02775, 1.17104)$, while the branch internal to the ellipse has its point deviations within $(\sqrt{t_1}, 19-5\sqrt{13}) \approx (0.86163, 0.97224)$. The worst cases points $\approx [17.83227, 2.47551]$ and $\approx [15.56302, 1.547248]$ giving maximal and minimal deviations respectively are marked in yellow in Fig. 5.2. The approximation error for the distance from any point of $K_1(x,y)=0$ to the ellipse is smaller than 20 %.

Similar computations for the curve $K_2(x,y)=0$ results in the deviation intervals $(1.50613, 1.88923)$ and $(2.38516, 2.73479)$ for the internal and the external branch respectively. The approximation error for the distance from the farthest point of the curve to the ellipse exceeds 35 %.

As for the curve $K_3(x,y)=0$ its deviation intervals are $(2.00277, 2.75171)$ and $(3.24829, 4.71782)$. The maximal error for the distance exceeds 55 %. □

To estimate the deviation of the curve (5.1) from the ellipse (3.5) in the general case, one needs to localize the real zeros of the equation (5.3).

Theorem 5.2. *For any specializations of parameters a, b and h satisfying the conditions of Theorem 5.1, equation (5.3) possesses* 4 *real zeros located in the intervals*

$$(-\infty; 0),\ (0, h),\ (h, \tilde{t}),\ (\tilde{t}, +\infty)\,; \quad \text{here}\ \tilde{t} := \frac{(a^2+b^2)^2}{a^2b^2}h\,.$$

Proof can be derived from the following inequalities:

$$\begin{array}{ll} \mathcal{K}_h(0) & \equiv -h^2a^4b^4[4\,h(b^2+a^2)^3+27\,a^4b^4] < 0, \\ \mathcal{K}_h(h) & \equiv h^3(a^2+b^2)^2[h(a^4+b^4+a^2b^2)^2+4\,a^4b^4(a^2+b^2)] > 0, \\ \mathcal{K}_h(\tilde{t}) & \equiv -\frac{h^2}{a^2b^2}[4\,(a^4+b^4)^3(a^2+b^2)^3+27\,a^6b^6(a^4+b^4+a^2b^2)^2] < 0, \end{array}$$

and $\mathcal{K}_h(-\infty) > 0, \mathcal{K}_h(+\infty) > 0$. It is also evident that $(a^2+b^2)^2/(a^2b^2) \geq 4$ for any nonzero specialization of parameters. This yields: $\tilde{t} > h$. □

We are interested in two least positive zeros of $\mathcal{K}_h(t)$ since they are responsible for deviations. Let us make some experiments on varying the parameters of the ellipse.

Example 5.2. *For $h = 1$, some approximate values for the least positive zeros $t_1, t_2, t_1 < t_2$ of $\mathcal{K}_h(t)$ under variations of parameters a and b are presented in the following table:*

a	800	400	200	200	250
b	24	12	6	5	6
t_1	0.62	0.45	0.28	0.20	0.17
t_2	1.81	3.35	9.09	17.18	26.01

One can watch that if the value $\tau := b/a$ remains fixed, then t_1 increases and t_2 decreases when both parameters parameters increase. If τ decreases while one of the parameters a or b remains fixed, then t_1 decreases and t_2 increases.

We now turn to the case of Sampson's approximation (4.7) for the point-to-ellipsoid distance in $\mathbb{R}^3$.

Theorem 5.3. *Let all the values a, b, c be distinct and $\sqrt{h} < \min\{a, b, c\}$. The values of the minimal and the maximal deviation of the surface*

$$K_{\sqrt{h}}(x, y, z) := G^2(x, y, z) - 4\,hS_4(x, y, z) = 0$$

from the ellipsoid (3.11) are attained at the cross-sections of this surface with the coordinate planes.

Proof. The starting point is similar to that in the proof of Theorem 5.1. Compose the Lagrange function

$$(x-\tilde{x})^2+(y-\tilde{y})^2+(z-\tilde{z})^2-\lambda_1 K_{\sqrt{h}}(x,y,z)-\lambda_2 G(\tilde{x},\tilde{y},\tilde{z})$$

and find its stationary points.

$$x-\tilde{x} = \frac{2\,\lambda_1 x}{a^2}\left(G(x,y,z)-\frac{2h}{a^2}\right), \tag{5.16}$$

$$y-\tilde{y} = \frac{2\,\lambda_1 y}{b^2}\left(G(x,y,z)-\frac{2h}{b^2}\right), \tag{5.17}$$

$$z-\tilde{z} = \frac{2\,\lambda_1 z}{c^2}\left(G(x,y,z)-\frac{2h}{c^2}\right), \tag{5.18}$$

$$x-\tilde{x}=-\lambda_2\tilde{x}/a^2,\ y-\tilde{y}=-\lambda_2\tilde{y}/b^2,\ z-\tilde{z}=-\lambda_2\tilde{z}/c^2, \tag{5.19}$$

$$K_{\sqrt{h}}(x,y,z)=0,\ G(\tilde{x},\tilde{y},\tilde{z})=0\,. \tag{5.20}$$

Rewrite the relations (5.19) in the form

$$x-\tilde{x}=\frac{\lambda_2 x}{\lambda_2-a^2},\ y-\tilde{y}=\frac{\lambda_2 y}{\lambda_2-b^2},\ z-\tilde{z}=\frac{\lambda_2 z}{\lambda_2-c^2}$$

and substitute into (5.16)–(5.18):

$$\begin{aligned}
\frac{\lambda_2 x}{\lambda_2-a^2} &= \frac{2\lambda_1 x}{a^2}\left(G(x,y,z)-\frac{2h}{a^2}\right),\\
\frac{\lambda_2 y}{\lambda_2-b^2} &= \frac{2\lambda_1 y}{b^2}\left(G(x,y,z)-\frac{2h}{b^2}\right),\\
\frac{\lambda_2 z}{\lambda_2-c^2} &= \frac{2\lambda_1 z}{c^2}\left(G(x,y,z)-\frac{2h}{c^2}\right).
\end{aligned}$$

If we assume that all the variables x,y,z are nonzero, then this system is equivalent to

$$\begin{aligned}
\frac{a^2\lambda_2}{\lambda_2-a^2} &= 2\lambda_1\left(G(x,y,z)-\frac{2h}{a^2}\right),\\
\frac{b^2\lambda_2}{\lambda_2-b^2} &= 2\lambda_1\left(G(x,y,z)-\frac{2h}{b^2}\right),\\
\frac{c^2\lambda_2}{\lambda_2-c^2} &= 2\lambda_1\left(G(x,y,z)-\frac{2h}{c^2}\right).
\end{aligned}$$

Elimination of λ_1 from this system yields

$$\lambda_2 G(x,y,z)+2\,h\left(1-\frac{\lambda_2}{a^2}-\frac{\lambda_2}{b^2}\right)=0,\ \lambda_2 G(x,y,z)+2\,h\left(1-\frac{\lambda_2}{a^2}-\frac{\lambda_2}{c^2}\right)=0$$

and this system is inconsistent provided that $b \neq c$. Therefore, for any stationary point of the constrained optimization problem, at least one of its coordinates x, y, z should be zero. This reduces the problem to the planar case considered in Theorem 5.1. □

We now turn to the case of the distance approximation given by formula (4.8). On restricting ourselves with the planar case, we intend to evaluate the deviation of the curve $d_{(2)}^2 = h$ from the ellipse (3.5). First represent this curve in the algebraic form

$$M_{\sqrt{h}}(x,y) := S_6(x,y)G^3(x,y) + 2\,G^2(x,y)S_4^2(x,y) - 8\,hS_4^3(x,y) = 0\,. \quad (5.21)$$

Here $S_4(x,y) := x^2/a^4 + y^2/b^4$, $S_6(x,y) := x^2/a^6 + y^2/b^6$. The curve (5.21) is of the 8th order and investigation of its behavior is more complicated in comparison of the 4th order curve (5.1).

Theorem 5.4. *Let $a > b > \sqrt{h}$. The values of the minimal and the maximal deviation of (5.21) from the ellipse (3.5) are either among the values*

$$\{|\sqrt{t_*} - a|,\ |\sqrt{t_{**}} - b|\} \quad (5.22)$$

where t_* and t_{**} stand for the positive zeros of the equations

$$3\,t^3-(8\,h+7\,a^2)t^2+5\,a^4t-a^6=0,\ 3\,t^3-(8\,h+7\,b^2)t^2+5\,b^4t-b^6=0 \quad (5.23)$$

respectively, or among the square roots of the positive zeros of a certain algebraic equation

$$\mathcal{M}_h(t) := \sum_{j=0}^{21} F_j t^{21-j} = 0 \quad (5.24)$$

of the degree 21 with respect to t and with coefficients $\{F_j\}_{j=0}^{21}$ polynomially dependent on the parameters a, b and h.

Proof is similar to that of Theorem 5.2. Equating the derivatives of the Lagrange function

$$(x-\tilde{x})^2 + (y-\tilde{y})^2 - \lambda_1 M_{\sqrt{h}}(x,y) - \lambda_2 G(\tilde{x},\tilde{y})$$

to zero one obtains the system of 6 equations. Two of them are[3]

$$
\begin{aligned}
x-\tilde{x} &= \frac{2\,\lambda_1 x}{a^2}\left(\frac{G^3}{a^4}+3\,G^2S_6+4\,GS_4^2+\frac{4}{a^2}\,G^2S_4-\frac{24\,h}{a^2}S_4^2\right), \qquad (5.25)\\
y-\tilde{y} &= \frac{2\,\lambda_1 y}{b^2}\left(\frac{G^3}{b^4}+3\,G^2S_6+4\,GS_4^2+\frac{4}{b^2}\,G^2S_4-\frac{24\,h}{b^2}S_4^2\right). \qquad (5.26)
\end{aligned}
$$

The third one coincides with (5.21) while the rest three coincide with (5.6), (5.7) and (5.9).

We follow on the further steps of the proof of Theorem 5.2. Equations (5.12) and (5.13) remain valid. Substitute next (5.11) into (5.25) and (5.26); the resulting equations can be split into the alternatives:

$$
x=0 \quad \text{or} \quad \frac{\lambda_2}{\lambda_2-a^2}=\frac{2\lambda_1}{a^2}\left(\frac{G^3}{a^4}+3\,G^2S_6+4\,GS_4^2+\frac{4}{a^2}\,G^2S_4-\frac{24\,h}{a^2}S_4^2\right);
$$

$$
y=0 \quad \text{or} \quad \frac{\lambda_2}{\lambda_2-b^2}=\frac{2\lambda_1}{b^2}\left(\frac{G^3}{b^4}+3\,G^2S_6+4\,GS_4^2+\frac{4}{b^2}\,G^2S_4-\frac{24\,h}{b^2}S_4^2\right).
$$

The first parts of these alternatives correspond to the values (5.22). From the second parts, it is possible to eliminate the parameter λ_1:

$$
\left(\frac{1}{\lambda_2}-\frac{1}{a^2}-\frac{1}{b^2}\right)(G^3(a^2+b^2)+4\,a^2b^2G^2S_4-24\,a^2b^2hS_4^2)
$$

$$
+G(G^2-3\,a^2b^2GS_6-4\,a^2b^2S_4^2)=0\,.
$$

Next, we substitute in this equation and in (5.21) the expressions (5.15) for x^2 and y^2. This results in the system of algebraic equations

$$
P_1(\lambda_2,t)=0,\ P_2(\lambda_2,t)=0\ ,\ \deg_{\lambda_2}P_1=11,\ \deg_{\lambda_2}P_2=14\,.
$$

Elimination of λ_2 from this system via the resultant computation yields (on excluding some extraneous factors) the equation (5.24) with

$$
\begin{aligned}
\mathcal{M}_h(t) \;:=\;& 144\,a^8b^8(3\,a^4+5\,a^2b^2+3\,b^4)(a^4+10\,a^2b^2+b^4)^4(a^2-b^2)^{10}t^{21}\\
&+24\,a^6b^6(a^4+10\,a^2b^2+b^4)^3(a^2-b^2)^8\Big[(3\,a^4+5\,a^2b^2+3\,b^4)\\
&\times[177\,(b^{12}+a^{12})+1826\,(b^{10}\,a^2+a^{10}b^2)-2273\,(b^8a^4+b^4a^8)-17316\,a^6b^6]h\\
&-18\,a^2b^2(a^2+b^2)(20\,(a^{12}+b^{12})+165\,(a^{10}b^2+a^2b^{10})-146\,(a^4b^8+a^8b^4)-430\,a^6b^6)\Big]t^{20}\\
&+\ldots
\end{aligned}
$$

[3]We skip the arguments x,y in the expressions for G,S_4 and S_6.

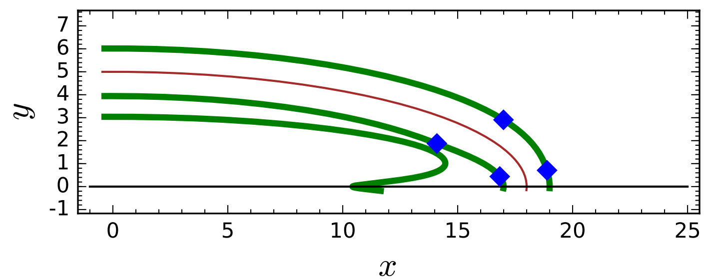


Figure 5.3: Curve $M_1(x,y)=0$ (green) in a vicinity of the ellipse $x^2/324+y^2/25=1$ (thin red line).

Complete expression can be found in [18]. □

Example 5.3. *Find the maximal and the minimal deviations of the curves* $\{M_j(x,y)=0\}_{j=1}^3$ *from the ellipse* $x^2/324+y^2/25=1$.

Solution. For the curve $M_1(x,y)=0$, polynomial (5.24) with coefficients of the orders up to 10^{98} possesses 11 positive zeros with the minimal ones

$$t_1 \approx 0.96136,\ t_2 \approx 0.98731,\ t_3 \approx 1.06466,\ t_4 \approx 1.34127\,.$$

The corresponding points in the curve $M_1(x,y)=0$ are as follows

$$X_1 \approx [16.83055, 0.44034],\ X_2 \approx [18.88723, 0.70845],$$

$$X_3 \approx [16.99212, 2.90776],\ X_4 \approx [14.09564, 1.87734]$$

(marked as boxes in Fig. 5.3). For the branch of the curve $M_1=0$ lying outside the ellipse, the deviation of any point is within the interval $(\sqrt{t_2},\sqrt{t_3}) \approx (0.99363, 1.03182)$. There are two branches of the curve lying inside the ellipse. The nearest to the ellipse has the deviation within the interval $(\sqrt{t_1},\sqrt{t_4}) \approx (0.98049, 1.15813)$.

The branch of the curve $M_2(x,y)=0$ lying inside the ellipse is not looking like an equidistant curve of the latter. Indeed, it even does not encircle the origin (Fig. 5.4). Thus, for the points lying inside the ellipse the validity of the distance approximation formula is limited to the set of points lying at the distances not exceeding 1.

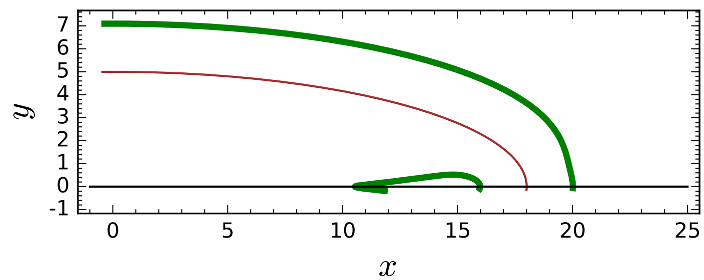


Figure 5.4: Curve $M_2(x, y) = 0$.

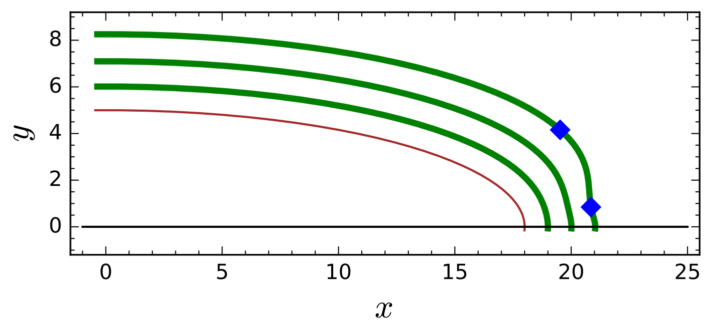


Figure 5.5: Branches of curves $\{M_j(x, y) = 0\}_{j=1}^{3}$ external to the ellipse.

Deviation of the branch of the curve $M_3(x, y) = 0$ lying outside the ellipse is within the interval $\approx (2.93246, 3.549460)$ with the coordinates of the worst case points being $\approx [20.84917, 0.84557]$ and $\approx [19.52559, 4.15596]$. The approximation error is within 20 %.

One can note that the behavior of the external to the ellipse branches of the curves $M_j(x, y) = 0$ are closer to the equidistant curve to this ellipse than their counterparts $K_j(x, y) = 0$ providing Sampson's distance.

Finally, let us make a comparison with the level curves provided by the Harker and O'Leary distance approximation given by (4.1). We transform the equations $d_{HO}^2 = h$ into algebraic form $L_{\sqrt{h}}(x, y) = 0$ using the coefficients of equation (3.7). The algebraic curve is of the same order like the curve $M_{\sqrt{h}}(x, y) = 0$, namely 8 and, for the ellipse of the present example, the plots of both curves in its vicinity look similar. However, a more careful

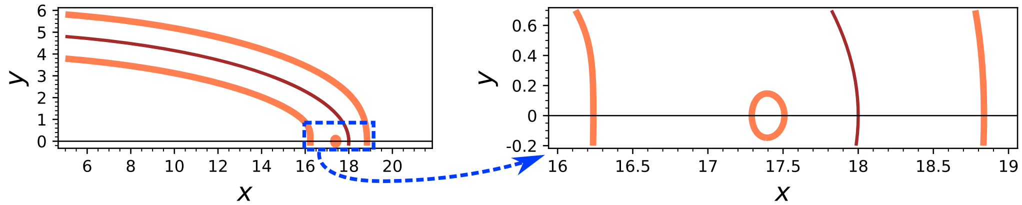


Figure 5.6: Curve $L_1(x, y) = 0$ (coral) in a vicinity of the ellipse.

investigation discovers a distinction in the qualitative behaviors. The curve $L_1(x, y) = 0$ possesses an extra tiny oval in between of the two concentric branches (Fig. 5.6). This oval surrounds one of the domains of inapplicability of the formula (3.7) located in Example 4.2 (Fig. 4.1).

The remained branches of the curve $L_1 = 0$ nearly coincide with those of $M_1 = 0$ with the worst possible deviation point at $\approx [16.237952, 0]$.

Deviation of the branch of the curve $L_2(x, y) = 0$ external to the ellipse displays increasing effect in comparison with that of $M_2(x, y) = 0$ with the worst possible point at $\approx [19.32363, 0]$ (i.e. 33 % error). It should be noted that the distinction in the plots of these curves can be watched only at the segments lying close to the major axis of the ellipse, i.e. for $x \in (19, 20)$. As for the branch of the curve $L_2(x, y) = 0$ internal to the ellipse, its deviation lies within $\approx (2, 4.57784)$ with the maximal error exceeding 100 %.

The plot of the curve $L_3(x, y) = 0$ displays the trend of drastic increasing of the error at the points close to x-axis. (Fig. 5.7). One can watch how the nearly everywhere parallel to the ellipse manifolds $L_2 = 0$ and $L_3 = 0$ gather in a point lying in this axis. This is also a mark of poor behavior for the approximation (3.7) already mentioned in Example 4.2. □

For the aim of further comparison of the distance approximations given by (4.1) and (4.8), we have generated several experiments on varying the parameters of the ellipse and the distance for its equidistant curves. The results of the qualitative analysis are as follows:

- For the case of a small eccentricity ellipse ($b \le a < \sqrt{2}b$), the both approximation formulas are nearly equivalent.

- For more elongated ellipses (i.e., with high eccentricity) the Harker-O'Leary distance approximation formula displays singularities in a vicinity of the major axis of the ellipse of the type similar to those discovered in Example 5.3.

- On ignoring these domains of outliers, it should be mentioned that the quality of approximation by the formula (4.1) for the points lying inside the ellipse is much lower compared with those lying outside.

- The last statement holds also for the approximation formula (4.8). Demonstrated in Fig. 5.4 degeneracy of the parallelism effect for the branch of $M_{\sqrt{h}} = 0$ internal to the ellipse is caused by its closeness to the curve bounding the domain of inapplicability of the formula (4.8) (i.e. the negativity of the involved radicand). We claim (without proof) that the maximal value of the parameter h for which an internal branch of the curve $M_{\sqrt{h}} = 0$ circles the origin is given as

  $$\widetilde{h} = \frac{-a^6 - 9\,a^4b^2 - 9\,a^2b^4 - b^6 + (a^4 + 4\,a^2b^2 + b^4)\sqrt{a^4 + 10\,a^2b^2 + b^4}}{16\,a^2b^2} .$$

  Thus, for $a = 40, b = 6$, one has $\widetilde{h} \approx 0.67101$ and therefore the internal to the ellipse branch of the curve $M_1 = 0$ does not circle the origin. The value $\sqrt{\widetilde{h}}$ can be treated as an upper bound for the applicability of the distance approximation formula (4.8) for the points internal to the ellipse. The asymptotical behavior for $\widetilde{h}$ as the b/a tends to zero is as follows: $\widetilde{h} \sim b^4/a^2$.

- On the contrary, the external to the ellipse branches of the curves $M_{\sqrt{h}} = 0$ look like more smoothly approximating the equidistant curves to this ellipse even for large values of h. This might be the reason for the conclusion that approximation (4.8) is more suitable for the applications concerned with collision detection of the objects rather than with the ellipse fitting ones.

We conclude the present section with an example demonstrating the applicability of the investigated approximations for the case of ellipsoid in $\mathbb{R}^4$.

Example 5.4. *For the ellipsoid in* $\mathbb{R}^4$ *given by*

$$X^{\top}\mathbf{A}X = 1 \quad \text{with} \quad \mathbf{A} = \frac{1}{127008}\left(\begin{array}{rrrr} 4523 & -230 & -230 & -3415 \\ -230 & 2318 & -1210 & -230 \\ -230 & -1210 & 2318 & -230 \\ -3415 & -230 & -230 & 4523 \end{array}\right)$$

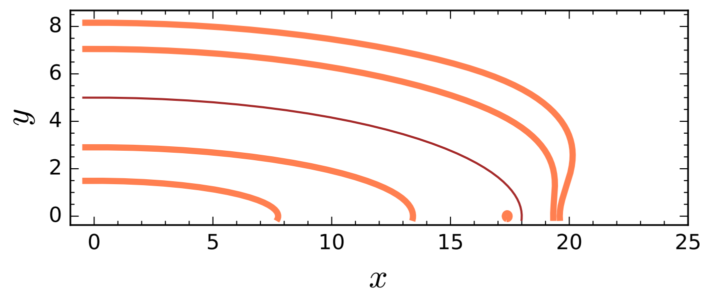


Figure 5.7: Curves $L_2(x, y) = 0$ and $L_3(x, y) = 0$ (coral) in a vicinity of the ellipse.

find approximations for the distance from some points lying in its axes at the distance $d = 1$.

Solution. We utilize Theorem 3.1 for computing the 16th order distance equation for specializations of the point X_0 coordinates; this permits us to evaluate the 4 dimensional counterpart for the approximation (4.1) as well. Approximations $d_{(1)}$ and $d_{(2)}$ are evaluated by formulas (4.11) and (4.12), respectively. The points in the following table lie in the axes of the ellipsoid

X_0	$\begin{bmatrix} \frac{5}{\sqrt{2}} \\ 0 \\ 0 \\ -\frac{5}{\sqrt{2}} \end{bmatrix}$	$\begin{bmatrix} \frac{3}{\sqrt{2}} \\ 0 \\ 0 \\ -\frac{3}{\sqrt{2}} \end{bmatrix}$	$\begin{bmatrix} 0 \\ \frac{7}{\sqrt{2}} \\ -\frac{7}{\sqrt{2}} \\ 0 \end{bmatrix}$	$\begin{bmatrix} 0 \\ \frac{5}{\sqrt{2}} \\ -\frac{5}{\sqrt{2}} \\ 0 \end{bmatrix}$	$\begin{bmatrix} 5 \\ -5 \\ -5 \\ 5 \end{bmatrix}$	$\begin{bmatrix} 4 \\ -4 \\ -4 \\ 4 \end{bmatrix}$	$\begin{bmatrix} \frac{15}{2} \\ \frac{15}{2} \\ \frac{15}{2} \\ \frac{15}{2} \end{bmatrix}$	$\begin{bmatrix} \frac{13}{2} \\ \frac{13}{2} \\ \frac{13}{2} \\ \frac{13}{2} \end{bmatrix}$
$G(X_0)$	0.56	−0.44	0.36	−0.31	0.23	−0.21	0.15	−0.14
d_{HO}	0.97	0.96	1.03	1.37	1.17	0.33	1.72	0.71
$d_{(1)}$	0.90	1.17	0.93	1.10	0.95	1.06	0.97	1.04
$d_{(2)}$	0.98	0.91	0.99	0.97	0.99	0.99	0.99	0.99

with corresponding values of its semi-principal axes equal to $4, 6, 9$ and 14. Distance equations computed for the last 4 points (lying in the longest axes of the ellipsoid) possesses double real zeros corresponding to imaginary points of the ellipsoid. For the points $X_0 = [4, -4, -4, 4]$ and $X_0 = [13/2, 13/2, 13/2, 13/2]$ the minimal positive zeros of the distance equations are lesser than 1. One can watch the increase of the error for the distance approximation d_{HO} due to the reasons mentioned in Section 4. □

6. Conclusion

We have treated the problem of finding an analytical approximation formulas for the point-to-ellipse or point-to-ellipsoid distance evaluation problems. The present paper can be considered as a continuation of the previous investigation by the authors [13]. At it is mentioned in Introduction, we aim at resolving the general distance equation presented in that paper. On executing this plan, we have not only rediscovered (surprisingly for us) the already known for several decades Sampson's approximation formula (1.2), but also found out its natural generalization by formula (1.3). Although the latter is a bit more complicated than the former, it often diminishes the approximation error.

It is challenging to discover the counterpart of the approximation (1.3) for an arbitrary algebraic manifold, and (even regardless the practical reasons, i.e. just for the sake of curiosity) to find the general form for the expansion (4.6) of the distance equation zero in powers of algebraic distance.

On the contrary, it is of practical interest to find a counterpart in $\mathbb{R}^n, n \geq 3$ of Theorem 5.4 for the distance approximation given by (1.3).

Results of Theorems 5.2 and 5.4 can be interpreted as related to the general problem of distance evaluation between two nonlinear algebraic curves (manifolds). Therefore, it seems sensible to seek for solution of this problem in terms of an appropriate distance equation construction. In [13] this equation is constructed for a pair of quadrics in $\mathbb{R}^n$. Unfortunately it is rather complicated (of the degree up to 12 for ellipses in $\mathbb{R}^2$ and up to 24 for ellipsoids in $\mathbb{R}^3$). We hardly expect the feasibility of extraction from its general form the explicit counterparts of the point-to-ellipsoid approximations from Theorem 4.4. A possible alternative approach has been proposed to the authors by one of the referees of the present paper: possessing the formulas for the distance approximations between a point and one of the quadrics represented in algebraic form, one can substitute for the coordinates of a point the parameterized representation of the second quadric. We hope to explore this pass and compare the results with the other approaches to the problem [19]. This problem is important in many applications such as robotics, CAD/CAM and computer graphics, and also in those arisen in Control Theory connected with the ellipsoidal form of the controlled object sets or the obstacles [20, 21].

Acknowledgements. The authors thank the referees for valuable suggestions that helped to significantly improve the quality of the paper.

References

[1] Y.H.Dai, Fast algorithms for projection on an ellipsoid. SIAM J.Optimiz. 16(4) (2006) 986–1006.

[2] N.Chernov, S.Wijewickrema, Algorithms for projecting points onto conics. J. Comput. Appl. Math. 251 (2013) 8–21.

[3] G.Sh. Tamasyan, A.A. Chumakov, Finding the distance between ellipsoids. Journal of Applied and Industrial Mathematics. 8 (3) (2014) 400–410.

[4] A. Best, S. Narang, D. Manocha, Real-time collision avoidance with elliptic agents. IEEE Intern. Conf. Robotics and Automation (ICRA 2016) (2016) 298–305.

[5] P.D. Sampson, Fitting conic sections to very scattered data: an iterative refinement of the Bookstein algorithm, Comput. Gr. Image Process. 18 (1982) 97–108.

[6] G. Taubin, Estimation of planar curves, surfaces and nonplanar space curves defined by implicit equations with application to edge and range image segmentation, IEEE Trans. Pattern Anal. Mach. Intell. 13 (11) (1991) 1115–1138.

[7] A.W. Fitzgibbon, M. Pilu, R.B. Fisher, Direct least-squares fitting of ellipses, IEEE Trans. Pattern Anal. Mach. Intell. 21 (5) (1999) 476–480.

[8] Z. Szpak, W. Chojnacki, A. van den Hengel, Guaranteed ellipse fitting with the Sampson distance, LNCS, Springer 7576 (1240) (2012) 87–100.

[9] P.L. Rosin, Analyzing error of fit function for ellipses. Pattern Recognition Lett. 17 (1999) 1461–1470.

[10] P.L. Rosin, Assessing error of fit function for ellipses. Graphical Models Image Process. 58 (1999) 494–502.

[11] P.L. Rosin, Evaluating Harker and O'Leary's distance approximation for ellipse fitting. Pattern Recognition Lett., vol. 28 (13) (2007) 1804–1807.

[12] A.Yu. Uteshev, M.V. Yashina, Distance computation from an ellipsoid to a linear or a quadric surface in $\mathbb{R}^n$, LNCS, Springer 4770 (2007) 392–401.

[13] A.Yu. Uteshev, M.V. Yashina, Metric problems for quadrics in multidimensional space, J.Symbolic Comput. 68 (1) (2015) 287–315.

[14] M. Harker, P. O'Leary, First order geometric distance (the myth of Sampsonus), Proc. of the British Machine Vision Conference (BMVC06), I, Edinburgh, UK (2006) 87–96.

[15] D.A. Cox, J. Little, D. O'Shea, Ideals, Varieties, and Algorithms, Springer, 2007.

[16] J.V. Uspensky, Theory of Equations, McGraw-Hill, 1948.

[17] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, 1986.

[18] A.Yu. Uteshev, Notebook. http://pmpu.ru/vf4/matricese/optimize/jcma1

[19] X.-D. Chen , J.-H. Yong, G.-Q. Zheng, J.-C. Paul, J.-G. Sun, Computing minimum distance between two implicit algebraic surfaces. CAD Comput. Aided Design, 38 (10) (2006) 1053–1061.

[20] A. Aleksandrov, Y. Chen, A. Platonov, L. Zhang, Stability analysis and uniform ultimate boundedness control synthesis for a class of nonlinear switched difference systems. J. Differ. Equ. Appl. 18 (9) (2012) 1545–1561.

[21] A.B. Kurzhanskij, On a team control problem under obstacles. Proc. Steklov Inst. Math. 291 (1) (2015) 128–142.